

Raw Sequence Listing
Patent Application US/07/762,762

05/21/92

15:15:10

#11
6/9/92

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: Thompson, Gregory A
Knauf, Vic C

(ii) TITLE OF INVENTION: Plant Desaturases-Compositions
and Uses

(iii) NUMBER OF SEQUENCES: 43

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: Calgene, Inc.

(B) STREET: 1920 Fifth Street

(C) CITY: Davis

(D) STATE: California

(E) COUNTRY: USA

(F) ZIP: 95616

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Diskette, 3.50 inch, 1.0 MB storage

(B) COMPUTER: Apple Macintosh

(C) OPERATING SYSTEM: Macintosh 6.0.7

(D) SOFTWARE: MicrosoftWord 4.0

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER: 07/762,762

(B) FILING DATE: 16-SEPT-1991

(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: PCT/US91/01746

(B) FILING DATE: 14-MAR-1991

(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:

Raw Sequence Listing
Patent Application US/07/762,762

05/21/92
15:15:13

54
55 (A) APPLICATION NUMBER: 07/615,784
56
57 (B) FILING DATE: 14-NOV-1990
58
59 (A) APPLICATION NUMBER: 07/567,373
60
61 (B) FILING DATE: 13-AUG-1990
62
63 (A) APPLICATION NUMBER: 07/494,106
64
65 (B) FILING DATE: 16-MAR-1990
66
67 (viii) ATTORNEY/AGENT INFORMATION:
68
69 (A) NAME: Lassen, Elizabeth
70
71 (B) REGISTRATION NUMBER: 31,845
72
73 (A) NAME: Donna E. Scherer
74
75 (B) REGISTRATION NUMBER: 34,719
76
77 (C) REFERENCE/DOCKET NUMBER: CGNE 69-4
78
79 (ix) TELECOMMUNICATION INFORMATION:
80
81 (A) TELEPHONE: (916) 753-6313
82
83 (B) TELEFAX: (916) 753-1510
84
85 (C) TELEX: 350370 CGNE
86
87
88 (2) INFORMATION FOR SEQ ID NO:1:
89
90 (i) SEQUENCE CHARACTERISTICS:
91
92 (A) LENGTH: 63 amino acids
93
94 (B) TYPE: amino acid
95
96 (D) TOPOLOGY: linear
97
98 (ii) MOLECULE TYPE: peptide
99
100
101 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:
102
103 Ala Ser Thr Leu Gly Ser Ser Thr Pro Lys Val Asp Asn Ala Lys Lys
104 1 5 10 15
105
106 Pro Phe Gln Pro Pro Arg Glu Val His Val Gln Val Thr His Xaa Met

Patent Application US/07/762,762

107 20 25 30
108
109 Pro Pro Gln Lys Ile Glu Ile Phe Lys Ser Ile Glu Gly Xaa Ala Glu
110 35 40 45
111
112 Gln Asn Ile Leu Val Xaa Leu Lys Pro Val Glu Lys Cys Trp Gln
113 50 55 60
114
115

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 56 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

131 Asp Phe Leu Pro Asp Pro Ala Xaa Glu Gly Phe Asp Glu Gln Val Lys
132 1 5 10 15
133
134 Glu Leu Arg Ala Arg Ala Lys Glu Ile Pro Asp Asp Tyr Phe Val Val
135 20 25 30
136
137 Leu Val Gly Asp Met Ile Thr Glu Glu Ala Leu Pro Thr Tyr Gln Thr
138 35 40 45
139
140 Met Leu Asn Thr Leu Asp Gly Val
141 50 55
142
143

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 14 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

158
159 Asp Glu Thr Gly Ala Ser Leu Thr Pro Trp Ala Val Trp Thr

Patent Application US/07/762,762

160 1 5 10
161
162
163 (2) INFORMATION FOR SEQ ID NO:4:
164
165 (i) SEQUENCE CHARACTERISTICS:
166
167 (A) LENGTH: 13 amino acids
168
169 (B) TYPE: amino acid
170
171 (D) TOPOLOGY: linear
172
173 (ii) MOLECULE TYPE: peptide
174
175
176 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:
177
178 Asp Leu Leu His Thr Tyr Leu Tyr Leu Ser Gly Arg Val
179 1 5 10
180
181
182 (2) INFORMATION FOR SEQ ID NO:5:
183
184 (i) SEQUENCE CHARACTERISTICS:
185
186 (A) LENGTH: 13 amino acids
187
188 (B) TYPE: amino acid
189
190 (D) TOPOLOGY: linear
191
192 (ii) MOLECULE TYPE: peptide
193
194
195 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:
196
197 Asp Met Arg Gln Ile Gln Lys Thr Ile Gln Tyr Leu Ile
198 1 5 10
199
200
201 (2) INFORMATION FOR SEQ ID NO:6:
202
203 (i) SEQUENCE CHARACTERISTICS:
204
205 (A) LENGTH: 17 amino acids
206
207 (B) TYPE: amino acid
208
209 (D) TOPOLOGY: linear
210
211 (ii) MOLECULE TYPE: peptide
212

Raw Sequence Listing
Patent Application US/07/762,762

05/21/92
15:15:20

213

214

215 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

216

217 Thr Glu Asn Ser Pro Tyr Leu Gly Phe Ile Tyr Thr Ser Phe Gln Glu

218 1 5 10 15

219

220 Arg

221

222

223 (2) INFORMATION FOR SEQ ID NO:7:

224

225 (i) SEQUENCE CHARACTERISTICS:

226

227 (A) LENGTH: 57 amino acids

228

229 (B) TYPE: amino acid

230

231 (D) TOPOLOGY: linear

232

233 (ii) MOLECULE TYPE: peptide

234

235

236 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

237

238 Asp Val Xaa Leu Ala Gln Ile Xaa Gly Thr Ile Ala Ser Asp Glu Lys

239 1 5 10 15

240

241 Arg His Glu Thr Ala Tyr Thr Lys Ile Val Glu Lys Leu Phe Glu Ile

242 20 25 30

243

244 Asp Pro Asp Gly Thr Val Leu Ala Phe Ala Asp Met Met Arg Lys Lys

245 35 40 45

246

247 Ile Xaa Met Pro Ala His Leu Met Tyr

248 50 55

249

250

251 (2) INFORMATION FOR SEQ ID NO:8:

252

253 (i) SEQUENCE CHARACTERISTICS:

254

255 (A) LENGTH: 4 amino acids

256

257 (B) TYPE: amino acid

258

259 (D) TOPOLOGY: linear

260

261 (ii) MOLECULE TYPE: peptide

262

263

264 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

265

Raw Sequence Listing
Patent Application US/07/762,762

05/21/92
15:15:27

266 Asp Asn Leu Phe

267 1

268

269 (2) INFORMATION FOR SEQ ID NO:9:

270

271 (i) SEQUENCE CHARACTERISTICS:

272

273 (A) LENGTH: 16 amino acids

274

275 (B) TYPE: amino acid

276

277 (D) TOPOLOGY: linear

278

279 (ii) MOLECULE TYPE: peptide

280

281

282 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

283

284 Xaa Xaa Phe Xaa Ala Val Xaa Gln Arg Leu Xaa Val Tyr Thr Ala Lys

285 1 5 10 15

286

287

288 (2) INFORMATION FOR SEQ ID NO:10:

289

290 (i) SEQUENCE CHARACTERISTICS:

291

292 (A) LENGTH: 14 amino acids

293

294 (B) TYPE: amino acid

295

296 (D) TOPOLOGY: linear

297

298 (ii) MOLECULE TYPE: peptide

299

300

301 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

302

303 Asp Tyr Ala Asp Ile Leu Glu Phe Leu Val Gly Arg Trp Lys

304 1 5 10

305

306

307 (2) INFORMATION FOR SEQ ID NO:11:

308

309 (i) SEQUENCE CHARACTERISTICS:

310

311 (A) LENGTH: 54 amino acids

312

313 (B) TYPE: amino acid

314

315 (D) TOPOLOGY: linear

316

317 (ii) MOLECULE TYPE: peptide

318

Raw Sequence Listing
Patent Application US/07/762,762

05/21/92
15:15:29

319

320

321 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

322

323

324 Val Ala Asp Leu Thr Gly Leu Ser Gly Glu Gly Arg Lys Ala Xaa Asp

325 1 5 10 15

326

327 Tyr Val Cys Gly Leu Pro Pro Arg Ile Arg Arg Leu Glu Glu Arg Ala

328 20 25 30

329

330 Gln Gly Arg Ala Lys Glu Gly Pro Val Val Pro Phe Ser Trp Ile Phe

331 35 40 45

332

333 Asp Arg Gln Val Lys Leu

334 50

335

336

337 (2) INFORMATION FOR SEQ ID NO:12:

338

339 (i) SEQUENCE CHARACTERISTICS:

340

341 (A) LENGTH: 1533 base pairs

342

343 (B) TYPE: nucleic acid

344

345 (C) STRANDEDNESS: double

346

347 (D) TOPOLOGY: linear

348

349 (ii) MOLECULE TYPE: cDNA to mRNA

350

351

352 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

353

354 GCTCACTTGT GTGGTGGAGG AGAAAAACAG AACTCACAAA AAGCTTTGCG ACTGCCAAGA 60

355

356 ACAACAACAA CAACAAGATC AAGAAGAAGA AGAAGAAGAT CAAAAATGGC TCTTCGAATC 120

357

358 ACTCCAGTGA CCTTGCAATC GGAGAGATAT CGTTCGTTTT CGTTTCCTAA GAAGGCTAAT 180

359

360 CTCAGATCTC CCAAATTCGC CATGGCCTCC ACCCTCGGAT CATCCACACC GAAGGTTGAC 240

361

362 AATGCCAAGA AGCCTTTTCA ACCTCCACGA GAGGTTTCATG TTCAGGTGAC GCACTCCATG 300

363

364 CCACCACAGA AGATAGAGAT TTTCAAATCC ATCGAGGGTT GGGCTGAGCA GAACATATTG 360

365

366 GTTCACCTAA AGCCAGTGGA GAAATGTTGG CAAGCACAGG ATTTCTTGCC GGACCCTGCA 420

367

368 TCTGAAGGAT TTGATGAACA AGTCAAGGAA CTAAGGGCAA GAGCAAAGGA GATTCTTGAT 480

369

370 GATTACTTTG TTGTTTTGGT TGGAGATATG ATTACAGAGG AAGCCCTACC TACTTACCAA 540

371

Patent Application US/07/762,762

372 ACAATGCTTA ATACCCTAGA TGGTGACGT GATGAGACTG GGGCTAGCCT TACGCCCTTG 600
373
374 GCTGTCTGGA CTAGGGCTTG GACAGCTGAA GAGAACAGGC ATGGCGATCT TCTCCACACC 660
375
376 TATCTCTACC TTTCTGGGCG GGTAGACATG AGGCAGATAC AGAAGACAAT TCAGTATCTC 720
377
378 ATTGGGTCAG GAATGGATCC TCGTACCGAA AACAGCCCCCT ACCTTGGGTT CATCTACACA 780
379
380 TCGTTTCAAG AGCGTGCCAC ATTTGTTTCT CACGGAAACA CCGCCAGGCA TGCAAAGGAT 840
381
382 CATGGGGACG TGAAACTGGC GCAAATTTGT GGTACAATCG CGTCTGACGA AAAGCGTCAC 900
383
384 GAGACCGCTT ATACAAAGAT AGTCGAAAAG CTATTCGAGA TCGATCCTGA TGGCACCGTT 960
385
386 CTTGCTTTTG CCGACATGAT GAGGAAAAAG ATCTCGATGC CCGCACACTT GATGTACGAT 1020
387
388 GGGCGTGATG ACAACCTCTT CGAACATTTT TCGGCGGTTG CCCAAAGACT CGGCGTCTAC 1080
389
390 ACCGCCAAAG ACTACGCCGA CATACTGGAA TTTCTGGTCG GCGGGTGGAA AGTGGCGGAT 1140
391
392 TTGACCGGCC TATCTGGTGA AGGGCGTAAA GCGCAAGATT ATGTTTGCGG GTTGCCACCA 1200
393
394 AGAATCAGAA GGCTGGAGGA GAGAGCTCAA GGGCGAGCAA AGGAAGGACC TGTGTTCCA 1260
395
396 TTCAGCTGGA TTTTCGATAG ACAGGTGAAG CTGTGAAGAA AAAAAAACG AGCAGTGAGT 1320
397
398 TCGGTTTCTG TTGGCTTATT GGGTAGAGGT TAAAACCTAT TTTAGATGTC TGTTCGTGT 1380
399
400 AATGTGGTTT TTTTCTTCT AATCTTGAAT CTGGTATTGT GTCGTTGAGT TCGCGTGTGT 1440
401
402 GTAAACTTGT GTGGCTGTGG ACATATTATA GAACTCGTTA TGCCAATTTT GATGACGGTG 1500
403
404 GTTATCGTCT CCCCTGGTGT TTTTATTG TTT 1533
405
406

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 396 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

Met Ala Leu Arg Ile Thr Pro Val Thr Leu Gln Ser Glu Arg Tyr Arg
-30 -25 -20

Patent Application US/07/762,762

425 Ser Phe Ser Phe Pro Lys Lys Ala Asn Leu Arg Ser Pro Lys Phe Ala
426 -15 -10 -5
427
428
429 Met Ala Ser Thr Leu Gly Ser Ser Thr Pro Lys Val Asp Asn Ala Lys
430 1 5 10 15
431
432 Lys Pro Phe Gln Pro Pro Arg Glu Val His Val Gln Val Thr His Ser
433 20 25 30
434
435 Met Pro Pro Gln Lys Ile Glu Ile Phe Lys Ser Ile Glu Gly Trp Ala
436 35 40 45
437
438 Glu Gln Asn Ile Leu Val His Leu Lys Pro Val Glu Lys Cys Trp Gln
439 50 55 60
440
441 Ala Gln Asp Phe Leu Pro Asp Pro Ala Ser Glu Gly Phe Asp Glu Gln
442 65 70 75
443
444 Val Lys Glu Leu Arg Ala Arg Ala Lys Glu Ile Pro Asp Asp Tyr Phe
445 80 85 90 95
446
447 Val Val Leu Val Gly Asp Met Ile Thr Glu Glu Ala Leu Pro Thr Tyr
448 100 105 110
449
450 Gln Thr Met Leu Asn Thr Leu Asp Gly Val Arg Asp Glu Thr Gly Ala
451 115 120 125
452
453 Ser Leu Thr Pro Trp Ala Val Trp Thr Arg Ala Trp Thr Ala Glu Glu
454 130 135 140
455
456 Asn Arg His Gly Asp Leu Leu His Thr Tyr Leu Tyr Leu Ser Gly Arg
457 145 150 155
458
459 Val Asp Met Arg Gln Ile Gln Lys Thr Ile Gln Tyr Leu Ile Gly Ser
460 160 165 170 175
461
462 Gly Met Asp Pro Arg Thr Glu Asn Ser Pro Tyr Leu Gly Phe Ile Tyr
463 180 185 190
464
465 Thr Ser Phe Gln Glu Arg Ala Thr Phe Val Ser His Gly Asn Thr Ala
466 195 200 205
467
468 Arg His Ala Lys Asp His Gly Asp Val Lys Leu Ala Gln Ile Cys Gly
469 210 215 220
470
471 Thr Ile Ala Ser Asp Glu Lys Arg His Glu Thr Ala Tyr Thr Lys Ile
472 225 230 235
473
474 Val Glu Lys Leu Phe Glu Ile Asp Pro Asp Gly Thr Val Leu Ala Phe
475 240 245 250 255
476
477 Ala Asp Met Met Arg Lys Lys Ile Ser Met Pro Ala His Leu Met Tyr

Patent Application US/07/762,762

```

478                260                265                270
479
480 Asp Gly Arg Asp Asp Asn Leu Phe Glu His Phe Ser Ala Val Ala Gln
481                275                280                285
482
483 Arg Leu Gly Val Tyr Thr Ala Lys Asp Tyr Ala Asp Ile Leu Glu Phe
484                290                295                300
485
486 Leu Val Gly Arg Trp Lys Val Ala Asp Leu Thr Gly Leu Ser Gly Glu
487                305                310                315
488
489 Gly Arg Lys Ala Gln Asp Tyr Val Cys Gly Leu Pro Pro Arg Ile Arg
490                320                325                330                335
491
492 Arg Leu Glu Glu Arg Ala Gln Gly Arg Ala Lys Glu Gly Pro Val Val
493                340                345                350
494
495 Pro Phe Ser Trp Ile Phe Asp Arg Gln Val Lys Leu
496                355                360
497
498

```

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 225 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

```

514
515 AAAAGAAAAA GGTAAGAAAA AAAACA ATG GCT CTC AAG CTC AAT CCT TTC CTT 53
516                MET Ala Leu Lys Leu Asn Pro Phe Leu
517                1                5
518
519 TCT CAA ACC CAA AAG TTA CCT TCT TTC GCT CTT CCA CCA ATG GCC AGT 101
520 Ser Gln Thr Gln Lys Leu Pro Ser Phe Ala Leu Pro Pro MET Ala Ser
521 10                15                20                25
522
523 ACC AGA TCT CCT AAG TTC TAC ATG GCC TCT ACC CTC AAG TCT GGT TCT 149
524 Thr Arg Ser Pro Lys Phe Tyr MET Ala Ser Thr Leu Lys Ser Gly Ser
525                30                35                40
526
527 AAG GAA GTT GAG AAT CTC AAG AAG CCT TTC ATG CCT CCT CGG GAG GTA 197
528 Lys Glu Val Glu Asn Leu Lys Lys Pro Phe MET Pro Pro Arg Glu Val
529                45                50                55
530

```

Patent Application US/07/762,762

531 CAT GTT CAG GTT ACC CAT TCT ATT GCC A 225
532 His Val Gln Val Thr His Ser Ile Ala
533 60 65
534
535

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1668 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

552
553 AAAAGAAAAA GGTAAGAAAA AAAACAATGG CTCTCAAGCT CAATCCTTTC CTTTCTCAAA 60
554
555 CCCAAAAGTT ACCTTCTTTC GCTCTTCCAC CAATGGCCAG TACCAGATCT CCTAAGTTCT 120
556
557 ACATGGCCTC TACCCTCAAG TCTGGTTCTA AGGAAGTTGA GAATCTCAAG AAGCCTTTCA 180
558
559 TGCCTCCTCG GGAGGTACAT GTTCAGGTTA CCCATTCTAT GCCACCCCAA AAGATTGAGA 240
560
561 TCTTTAAATC CCTAGACAAT TGGGCTGAGG AGAACATTCT GGTTCATCTG AAGCCAGTTG 300
562
563 AGAAATGTTG GCAACCGCAG GATTTTTTGC CAGATCCCGC CTCTGATGGA TTTGATGAGC 360
564
565 AAGTCAGGGA ACTCAGGGAG AGAGCAAAGG AGATTCTTGA TGATTATTTT GTTGTTTTGG 420
566
567 TTGGAGACAT GATAACGGAA GAAGCCCTTC CCACTTATCA AACAATGCTG AATACCTTGG 480
568
569 ATGGAGTTCT GGATGAAACA GGTGCAAGTC CTACTTCTTG GGCAATTGGG ACAAGGGCAT 540
570
571 GGA CTGCGGA AGAGAATAGA CATGGTGACC TCCTCAATAA GTATCTCTAC CTATCTGGAC 600
572
573 GAGTGGACAT GAGGCAAATT GAGAAGACAA TTCAATATTT GATTGGTTCA GGAATGGATC 660
574
575 CACGGACAGA AAACAGTCCA TACCTTGGGT TCATCTATAC ATCATTCCAG GAAAGGGCAA 720
576
577 CCTTCATTTT TCATGGGAAC ACTGCCCCGAC AAGCCAAAGA GCATGGAGAC ATAAAGTTGG 780
578
579 CTCAAATATG TGGTACAATT GCTGCAGATG AGAAGCGCCA TGAGACAGCC TACACAAAGA 840
580
581 TAGTGGAAAA ACTCTTTGAG ATTGATCCTG ATGGAAGTGT TTTGGCTTTT GCTGATATGA 900
582
583 TGAGAAAGAA AATTTCTATG CCTGCACACT TGATGTATGA TGGCCGAGAT GATAATCTTT 960

Patent Application US/07/762,762

584
585 TTGACCACTT TTCAGCTGTT GCGCAGCGTC TTGGAGTCTA CACAGCAAAG GATTATGCAG 1020
586
587 ATATATTGGA GTTCTTGGTG GGCAGATGGA AGGTGGATAA ACTAACGGGC CTTTCAGCTG 1080
588
589 AGGGACAAAA GGCTCAGGAC TATGTTTGTC GGTACCTCC AAGAATTAGA AGGCTGGAAG 1140
590
591 AGAGAGCTCA AGGAAGGGCA AAGGAAGCAC CCACCATGCC TTTCAGCTGG ATTTTCGATA 1200
592
593 GGCAAGTGAA GCTGTAGGTG GCTAAAGTGC AGGACGAAAC CGAAATGGTT AGTTTCACTC 1260
594
595 TTTTTCATGC CCATCCCTGC AGAATCAGAA GTAGAGGTAG AATTTTGTAG TTGCTTTTTT 1320
596
597 ATTACAAGTC CAGTTTAGTT TAAGGTCTGT GGAAGGGAGT TAGTTGAGGA GTGAATTTAG 1380
598
599 TAAGTTGTAG ATACAGTTGT TTCTTGTGTT GTCATGAGTA TGCTGATAGA GAGCAGCTGT 1440
600
601 AGTTTTGTTG TTGTGTTCTT TTATATGGTC TCTTGATGA GTTTCCTTTC TTTCTTTTC 1500
602
603 TTCTTTCCTT TCCTCTCTCT CTCTCTCTCT CTCTCTCTTT TTCTCTTATC CCAAGTGTCT 1560
604
605 CAAGTATAAT AAGCAAACGA TCCATGTGGC AATTTTGATG ATGGTGATCA GTCTCACAAC 1620
606
607 TTGATCTTTT GTCTTCTATT GGAAACACAG CCTGCTTGTT TGAAAAAA 1668
608
609

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 396 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

624
625 MET Ala Leu Lys Leu Asn Pro Phe Leu Ser Gln Thr Gln Lys Leu Pro
626 1 5 10 15
627
628 Ser Phe Ala Leu Pro Pro MET Ala Ser Thr Arg Ser Pro Lys Phe Tyr
629 20 25 30
630
631 MET Ala Ser Thr Leu Lys Ser Gly Ser Lys Glu Val Glu Asn Leu Lys
632 35 40 45
633
634 Lys Pro Phe MET Pro Pro Arg Glu Val His Val Gln Val Thr His Ser
635 50 55 60
636

Patent Application US/07/762,762

```

637 MET Pro Pro Gln Lys Ile Glu Ile Phe Lys Ser Leu Asp Asn Trp Ala
638 65 70 75 80
639
640 Glu Glu Asn Ile Leu Val His Leu Lys Pro Val Glu Lys Cys Trp Gln
641 85 90 95
642
643
644 Pro Gln Asp Phe Leu Pro Asp Pro Ala Ser Asp Gly Phe Asp Glu Gln
645 100 105 110
646
647 Val Arg Glu Leu Arg Glu Arg Ala Lys Glu Ile Pro Asp Asp Tyr Phe
648 115 120 125
649
650 Val Val Leu Val Gly Asp MET Ile Thr Glu Glu Ala Leu Pro Thr Tyr
651 130 135 140
652
653 Gln Thr MET Leu Asn Thr Leu Asp Gly Val Arg Asp Glu Thr Gly Ala
654 145 150 155 160
655
656 Ser Pro Thr Ser Trp Ala Ile Trp Thr Arg Ala Trp Thr Ala Glu Glu
657 165 170 175
658
659 Asn Arg His Gly Asp Leu Leu Asn Lys Tyr Leu Tyr Leu Ser Gly Arg
660 180 185 190
661
662 Val Asp MET Arg Gln Ile Glu Lys Thr Ile Gln Tyr Leu Ile Gly Ser
663 195 200 205
664
665 Gly MET Asp Pro Arg Thr Glu Asn Ser Pro Tyr Leu Gly Phe Ile Tyr
666 210 215 220
667
668 Thr Ser Phe Gln Glu Arg Ala Thr Phe Ile Ser His Gly Asn Thr Ala
669 225 230 235 240
670
671 Arg Gln Ala Lys Glu His Gly Asp Ile Lys Leu Ala Gln Ile Cys Gly
672 245 250 255
673
674 Thr Ile Ala Ala Asp Glu Lys Arg His Glu Thr Ala Tyr Thr Lys Ile
675 260 265 270
676
677 Val Glu Lys Leu Phe Glu Ile Asp Pro Asp Gly Thr Val Leu Ala Phe
678 275 280 285
679
680 Ala Asp MET MET Arg Lys Lys Ile Ser MET Pro Ala His Leu MET Tyr
681 290 295 300
682
683 Asp Gly Arg Asp Asp Asn Leu Phe Asp His Phe Ser Ala Val Ala Gln
684 305 310 315 320
685
686 Arg Leu Gly Val Tyr Thr Ala Lys Asp Tyr Ala Asp Ile Leu Glu Phe
687 325 330 335
688
689 Leu Val Gly Arg Trp Lys Val Asp Lys Leu Thr Gly Leu Ser Ala Glu

```

Raw Sequence Listing
Patent Application US/07/762,762

05/21/92
15:16:19

690		340	345	350
691				
692	Gly Gln Lys Ala Gln Asp Tyr Val Cys Arg Leu Pro Pro Arg Ile Arg			
693		355	360	365
694				
695	Arg Leu Glu Glu Arg Ala Gln Gly Arg Ala Lys Glu Ala Pro Thr MET			
696		370	375	380
697				
698	Pro Phe Ser Trp Ile Phe Asp Arg Gln Val Lys Leu			
699		385	390	395
700				
701				

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 117 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

719	TGAGAGATAG TGTGAGAGCA TTAGCCTTAG AGAGAGAGAG AGAGAGCTTG TGTCTGAAAG	60
720		
721	AATCCACAA ATG GCA TTG AAG CTT AAC CCT TTG GCA TCT CAG CCT TAC AAC	111
722	MET Ala Leu Lys Leu Asn Pro Leu Ala Ser Gln Pro Tyr Asn	
723	1 5 10	
724		
725	TTC CCT	117
726	Phe Pro	
727	15	
728		
729		

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 176 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

Patent Application US/07/762,762

743

744

745 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

746

747 ACT TCA TGG GCT ATT TGG ACA AGA GCT TGG ACT GCA GAA GAG AAC CGA 48

748 Thr Ser Trp Ala Ile Trp Thr Arg Ala Trp Thr Ala Glu Glu Asn Arg

749 1 5 10 15

750

751

752 CAC GGT GAT CTT CTC AAT AAG TAT CTT TAC TTG TCT GGA CGT GTT GAC 96

753 His Gly Asp Leu Leu Asn Lys Tyr Leu Tyr Leu Ser Gly Arg Val Asp

754 20 25 30

755

756 ATG AGG CAG ATT GAA AAG ACC ATT CAG TAC TTG ATT GGT TCT GGA ATG 144

757 MET Arg Gln Ile Glu Lys Thr Ile Gln Tyr Leu Ile Gly Ser Gly MET

758 35 40 45

759

760 GAT CCT AGA ACA GAG AAC AAT CCT TAC CTC GG 176

761 Asp Pro Arg Thr Glu Asn Asn Pro Tyr Leu

762 50 55

763

764

765 (2) INFORMATION FOR SEQ ID NO:19:

766

767 (i) SEQUENCE CHARACTERISTICS:

768

769 (A) LENGTH: 1495 base pairs

770

771 (B) TYPE: nucleic acid

772

773 (C) STRANDEDNESS: double

774

775 (D) TOPOLOGY: linear

776

777 (ii) MOLECULE TYPE: cDNA to mRNA

778

779

780 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

781

782 TGAGAGATAG TGTGAGAGCA TTAGCCTTAG AGAGAGAGAG AGAGAGCTTG TGTCTGAAAG 60

783

784 AATCCACAAA TGGCATTGAA GCTTAACCTT TTGGCATCTC AGCCTTACAA CTTCCCTTCC 120

785

786 TCGGCTCGTC CGCCAATCTC TACTTTCAGA TCTCCCAAGT TCCTCTGCCT CGCTTCTTCT 180

787

788 TCTCCCGCTC TCAGCTCCAA GGAGGTTGAG AGTTTGAAGA AGCCATTAC ACCACCTAAG 240

789

790 GAAGTGCACG TTCAAGTCCT GCATTCCATG CCACCCCAAG AGATCGAGAT CTTCAAATCC 300

791

792 ATGGAAGACT GGGCCGAGCA GAACCTTCTA ACTCAGCTCA AAGACGTGGA GAAGTCGTGG 360

793

794 CAGCCCCAGG ACTTCTTACC CGACCCTGCA TCCGATGGGT TCGAAGATCA GGTTAGAGAG 420

795

Patent Application US/07/762,762

796 CTAAGAGAGA GGGCAAGAGA GCTCCCTGAT GATTACTTCG TTGTTCTGGT GGGAGACATG 480
797
798 ATCACGGAAG AGGCGCTTCC GACCTATCAA ACCATGTTGA ACACTTTGGA TGGAGTGAGG 540
799
800 GATGAAACTG GCGCTAGCCC CACTTCATGG GCTATTTGGA CAAGAGCTTG GACTGCAGAA 600
801
802 GAGAACCGAC ACGGTGATCT TCTCAATAAG TATCTTTACT TGTCTGGACG TGTTGACATG 660
803
804 AGGCAGATTG AAAAGACCAT TCAGTACTTG ATTGGTTCTG GAATGGATCC TAGAACAGAG 720
805
806 AACAATCCTT ACCTCGGCTT CATCTACACT TCATTCCAAG AAAGAGCCAC CTTCATCTCT 780
807
808 CACGGAAACA CAGCTCGCCA AGCCAAAGAG CACGGAGACC TCAAGCTAGC CCAAATCTGC 840
809
810 GGCACAATAG CTGCAGACGA GAAGCGTCAT GAGACAGCTT ACACCAAGAT AGTTGAGAAG 900
811
812 CTCTTTGAGA TTGATCCTGA TGGTACTGTG ATGGCGTTTG CAGACATGAT GAGGAAGAAA 960
813
814 ATCTCGATGC CTGCTCACTT GATGTACGAT GGGCGGGATG AAAGCCTCTT TGACAACTTC 1020
815
816 TCTTCTGTTG CTCAGAGGCT CGGTGTTTAC ACTGCCAAAG ACTATGCCGA CATTCTTGAG 1080
817
818 TTTTGTGTTG GGAGGTGGAA GATTGAGAGC TTGACCGGGC TTCAGGTGA AGGAAACAAA 1140
819
820 GCGCAAGAGT ACTTGTGTGG GTTGA CTCCA AGAATCAGGA GGTGGATGA GAGAGCTCAA 1200
821
822 GCAAGAGCCA AGAAAGGACC CAAGGTTTCCT TTCAGCTGGA TACATGACAG AGAAGTGCAG 1260
823
824 CTCTAAAAAG GAACAAAGCT ATGAAACCTT TTCACTCTCC GTCGTCCCTC ATTTGATCTA 1320
825
826 TCTGCTCTTG AAATTGGTGT AGATTACTAT GGTTTGTGAT ATTGTTCTGT GGTCTAGTTA 1380
827
828 CAAAGTTGAG AAGCAGTGAT TTAGTAGCTT TGTGTTTCC AGTCTTTAAA TGTMTTGTG 1440
829
830 TTTGGTCCTT TTAGTAAACT TGTGTAGTT AAATCAGTTG AACTGTTTGG TCTGT 1495
831
832

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 398 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

MET Ala Leu Lys Leu Asn Pro Leu Ala Ser Gln Pro Tyr Asn Phe Pro

Patent Application US/07/762,762

849	1	5	10	15
850				
851	Ser Ser Ala Arg Pro Pro Ile Ser Thr Phe Arg Ser Pro Lys Phe Leu			
852	20	25	30	
853				
854	Cys Leu Ala Ser Ser Ser Pro Ala Leu Ser Ser Lys Glu Val Glu Ser			
855	35	40	45	
856				
857	Leu Lys Lys Pro Phe Thr Pro Pro Lys Glu Val His Val Gln Val Leu			
858	50	55	60	
859				
860	His Ser MET Pro Pro Gln Lys Ile Glu Ile Phe Lys Ser MET Glu Asp			
861	65	70	75	80
862				
863	Trp Ala Glu Gln Asn Leu Leu Thr Gln Leu Lys Asp Val Glu Lys Ser			
864	85	90	95	
865				
866	Trp Gln Pro Gln Asp Phe Leu Pro Asp Pro Ala Ser Asp Gly Phe Glu			
867	100	105	110	
868				
869	Asp Gln Val Arg Glu Leu Arg Glu Arg Ala Arg Glu Leu Pro Asp Asp			
870	115	120	125	
871				
872	Tyr Phe Val Val Leu Val Gly Asp MET Ile Thr Glu Glu Ala Leu Pro			
873	130	135	140	
874				
875	Thr Tyr Gln Thr MET Leu Asn Thr Leu Asp Gly Val Arg Asp Glu Thr			
876	145	150	155	160
877				
878	Gly Ala Ser Pro Thr Ser Trp Ala Ile Trp Thr Arg Ala Trp Thr Ala			
879	165	170	175	
880				
881	Glu Glu Asn Arg His Gly Asp Leu Leu Asn Lys Tyr Leu Tyr Leu Ser			
882	180	185	190	
883				
884	Gly Arg Val Asp MET Arg Gln Ile Glu Lys Thr Ile Gln Tyr Leu Ile			
885	195	200	205	
886				
887	Gly Ser Gly MET Asp Pro Arg Thr Glu Asn Asn Pro Tyr Leu Gly Phe			
888	210	215	220	
889				
890	Ile Tyr Thr Ser Phe Gln Glu Arg Ala Thr Phe Ile Ser His Gly Asn			
891	225	230	235	240
892				
893	Thr Ala Arg Gln Ala Lys Glu His Gly Asp Leu Lys Leu Ala Gln Ile			
894	245	250	255	
895				
896	Cys Gly Thr Ile Ala Ala Asp Glu Lys Arg His Glu Thr Ala Tyr Thr			
897	260	265	270	
898				
899	Lys Ile Val Glu Lys Leu Phe Glu Ile Asp Pro Asp Gly Thr Val Met			
900	275	280	285	
901				

Patent Application US/07/762,762

902 Ala Phe Ala Asp MET MET Arg Lys Lys Ile Ser Met Pro Ala His Leu
903 290 295 300
904
905 Met Tyr Asp Gly Arg Asp Glu Ser Leu Phe Asp Asn Phe Ser Ser Val
906 305 310 315 320
907
908 Ala Gln Arg Leu Gly Val Tyr Thr Ala Lys Asp Tyr Ala Asp Ile Leu
909 325 330 335
910
911 Glu Phe Leu Val Gly Arg Trp Lys Ile Glu Ser Leu Thr Gly Leu Ser
912 340 345 350
913
914 Gly Glu Gly Asn Lys Ala Gln Glu Tyr Leu Cys Gly Leu Thr Pro Arg
915 355 360 365
916
917 Ile Arg Arg Leu Asp Glu Arg Ala Gln Ala Arg Ala Lys Lys Gly Pro
918 370 375 380
919
920 Lys Val Pro Phe Ser Trp Ile His Asp Arg Glu Val Gln Leu
921 385 390 395
922
923

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 29 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: synthetic oligonucleotide mixture

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

GCTAAGCTTA ARGARATHCC AGAYGAYTA 29

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 29 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

Raw Sequence Listing
Patent Application US/07/762,762

05/21/92
15:16:56

955 (D) TOPOLOGY: linear
956
957 (ii) MOLECULE TYPE: other nucleic acid
958
959 (A) DESCRIPTION: synthetic oligonucleotide mixture
960
961 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:
962
963 GCTAAGCTTA ARGARATHCC GGAYGAYTA 29
964
965
966
967 (2) INFORMATION FOR SEQ ID NO:23:
968
969 (i) SEQUENCE CHARACTERISTICS:
970
971 (A) LENGTH: 29 base pairs
972
973 (B) TYPE: nucleic acid
974
975 (C) STRANDEDNESS: single
976
977 (D) TOPOLOGY: linear
978
979 (ii) MOLECULE TYPE: other nucleic acid
980
981 (A) DESCRIPTION: synthetic oligonucleotide mixture
982
983 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:
984
985 GCTAAGCTTA ARGARATHCC CGAYGAYTA 29
986
987
988 (2) INFORMATION FOR SEQ ID NO:24:
989
990 (i) SEQUENCE CHARACTERISTICS:
991
992 (A) LENGTH: 29 base pairs
993
994 (B) TYPE: nucleic acid
995
996 (C) STRANDEDNESS: single
997
998 (D) TOPOLOGY: linear
999
1000 (ii) MOLECULE TYPE: other nucleic acid
1001
1002 (A) DESCRIPTION: synthetic oligonucleotide mixture
1003
1004 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:
1005
1006 GCTAAGCTTA ARGARATHCC TGAYGAYTA 29
1007

Raw Sequence Listing
Patent Application US/07/762,762

05/21/92
15:17:04

1008
1009 (2) INFORMATION FOR SEQ ID NO:25:
1010
1011 (i) SEQUENCE CHARACTERISTICS:
1012
1013 (A) LENGTH: 26 base pairs
1014
1015 (B) TYPE: nucleic acid
1016
1017 (C) STRANDEDNESS: single
1018
1019 (D) TOPOLOGY: linear
1020
1021 (ii) MOLECULE TYPE: other nucleic acid
1022
1023 (A) DESCRIPTION: synthetic oligonucleotide mixture
1024
1025 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:
1026
1027 AGCGAATTCG TRTTNAGCAT NGTYTG 26
1028
1029
1030 (2) INFORMATION FOR SEQ ID NO:26:
1031
1032 (i) SEQUENCE CHARACTERISTICS:
1033
1034 (A) LENGTH: 26 base pairs
1035
1036 (B) TYPE: nucleic acid
1037
1038 (C) STRANDEDNESS: single
1039
1040 (D) TOPOLOGY: linear
1041
1042 (ii) MOLECULE TYPE: other nucleic acid
1043
1044 (A) DESCRIPTION: synthetic oligonucleotide mixture
1045
1046 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:
1047
1048 AGCGAATTCG TRTTYAACAT NGTYTG 26
1049
1050
1051 (2) INFORMATION FOR SEQ ID NO:27:
1052
1053 (i) SEQUENCE CHARACTERISTICS:
1054
1055 (A) LENGTH: 3440 base pairs
1056
1057 (B) TYPE: nucleic acid
1058
1059 (C) STRANDEDNESS: double
1060

Raw Sequence Listing
Patent Application US/07/762,762

05/21/92
15:17:11

1061 (D) TOPOLOGY: linear

1062

1063 (ii) MOLECULE TYPE: genomic DNA

1064

1065

1066 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

1067

1068 TCTAGAATTC TCTAATTACG TCTGTTTGTT CTATTTTTTA TATGATATCA AATATTCGTC 60

1069

1070 ATAAATATAT GGTTTAAGAT GCCAAAAAAT TATTTACTTG GTGAATATAA TACGTTAAAT 120

1071

1072 ATTAGAAATA CATCATTTAG TTAAATAAAT AACCAAAAAC CAAAAATTCA TATCCGCGCT 180

1073

1074 GGC GCGCGGT CAGGGTCTCG TTAGTTTTAA AATCAATGCA GTTTACAATT AATTTCCAGC 240

1075

1076 TGAAAATAAG TATAATTTGT ATTGAAATTA TAAAGTGACA TTTTTTGTGT AACAAATATT 300

1077

1078 TTGTGTAACA AGAATTAAAA AAAAAACAG AAAATACTCA GCTTTTTTAA TAATAAAAAA 360

1079

1080 AATTAATTGA GTTAGAAAAT TGTGTACCA ATAACAAAAG ATTTATATGG AATTATAAAA 420

1081

1082 TCAACACACC AATAACACAA GACTTTTTAA AAATTTAAGA ATAATATAAG CAATAACAAT 480

1083

1084 AGAATCTTCA AATTCTTCAA ATCCTTAAAA ATCAATCTCC CACTATTAAT CCCCTTAGT 540

1085

1086 TTTAGTTGGT AATGGCAACG TTTGTTGACT ACCGTATTGT AACTTTTGTC AAATTGTCAT 600

1087

1088 AAATACGTGT CAAACTCTGG TAAAAAATTA GTCTGCTACA TCTGTCTTTT ATTTATAAAA 660

1089

1090 CACAGCTGTT AATCAGAATT TGGTTTATTA AATCAACAAC CTGCACGAAA CTTGTGTGAG 720

1091

1092 CATATTTTGT CTGTTTCTGG TTCATGACCT TCTTCCGCAT GATGGCCAAG TGTAATGGCC 780

1093

1094 ACTTGCAAGA GCGTTTCTTC AACGAGATAA GTCGAACAAA TATTTGTCCG TTACGACCAC 840

1095

1096 ATATAAAATC TCCCATCTC TATATATAAT ACCAGCATTC ACCATCATGA ATACCTCAAA 900

1097

1098 TCCAATCTC ACAAATACTT CAATAAAAAG ACCAAAAAAA ATTAAAGCAA AGAAAAGCCT 960

1099

1100 TCTTGTGCAC AAAAAAAAAA GAAGCCTTCT AGGTTTTCAC GACATGAAGT TCACTACTCT 1020

1101

1102 AATGGTCATC ACATTGGTGA TAATCGCCAT CTCGTCTCCT GTTCCAATTA GAGCAACCAC 1080

1103

1104 GGTGAAAGT TTCGGAGAAG TGGCACAATC GTGTGTTGTG ACAGAACTCG CCCCATGCTT 1140

1105

1106 ACCAGCAATG ACCACGGCAG GAGACCCGAC TACAGAATGC TGCGACAAAC TGGTAGAGCA 1200

1107

1108 GAAACCATGT CTTTGTGGTT ATATTCGAAA CCCAGCCTAT AGTATGTATG TTACTTCTCC 1260

1109

1110 AAACGGTCGC AAAGTCTTAG ATTTTGTGTA GGTTCCTTTT CCTAGTTGTT AAATCTCTCA 1320

1111

1112 AGACATTGCT AAGAAAAATA TTATTAAAAA TAAAAGAATC AAAC TAGATC TGATGTAACA 1380

1113

Raw Sequence Listing
Patent Application US/07/762,76205/21/92
15:17:19

1114 ATGAATCATC ATGTTATGGT TGAAGCTTAT ATGCTGAAGT GTTTGATTTT ATATATGTGT 1440
1115
1116 GTGTGTGTGT CCTGCTCAAT TTTTGAAACA CACACGTTTC TCCTGATTTG GATTTAATTT 1500
1117
1118 ATATTTTGAG TTAAAAAAGAA GAAAAAGATG GAATGCTATT TATACAAGTT GATGAAAAAG 1560
1119
1120 TGGAAGTACA ATTTAGATAT CTCCTACACT TAAAGAATGA AACAAATAATA GACTTACGAA 1620
1121
1122 ACAAATGAAA AATACATAAA TTGTGACAAA TCAACGTCCG ATGACGAGTT TATTATTAAA 1680
1123
1124 AATTTGTGTG AAGGACTAGC AGTTCAACCA AATGATATTG AACATATACA TCAACAAAATA 1740
1125
1126 TGATAATCAT AAAAGAGAGA ATGGGGGGGG GGTGTCGTTT ACCAGAAACC TCTTTTCTC 1800
1127
1128 AGCTCGCTAA AACCTACCA CTAGAGACCT AGCTCTGACC GTCGGCTCAT CGGTGCCGGA 1860
1129
1130 GGTGTAACCT TTCTTTCCCA TGACCCGAAA CCTCTCTTTC CCAACTCACG AAAACCCTAC 1920
1131
1132 AATCAAAAAC CTAGCTCCGA CCGTCGGCTC ATCGGTGCCG AAGGTGTAAC CTTTCTCTCC 1980
1133
1134 CATCATAGTT TCTCGTAAAT GAAAGCTAAT TGGGCAATCG ATTTTTTAAT GTTTAAACCA 2040
1135
1136 TGCCAAGCCA TTTCTTATAG GACAATTGTC AATAATAGCA TCTTTTGAGT TTTGTCTCAA 2100
1137
1138 AAGTGACACT AGAAGAAAAA AGTCACAAAA ATGACATTCA TTAAAAAGTA AAATATCCCT 2160
1139
1140 AATACCTTTG GTTTAAATTA AATAAGTAAA CAAAAATAAA TAAAAACAAA TAAAAATAAA 2220
1141
1142 ATAAAAAATG AAAAAAAGAA ATTTTTTTTAT AGTTTCAGAT TATATGTTTT CAGATTCGAA 2280
1143
1144 ATTTTTTAAA TTCCCTTTTT TAAATTTTCT TTTTGAAT TTTTTTTTTT GAAATTTTTT 2340
1145
1146 GAAACTGTTT TTAAATTTT TATTTTAAAT TTTTAGTAT TTATTTTTTA TTTTATAAAA 2400
1147
1148 TTTTAAACGC TAATTCCAAA ACTCCCCCCC CCCCCCCCCC CCAATTCTC TCCTAGTCTT 2460
1149
1150 TTTCTCTTTC TTATATTTGG GCTTCTATCT TCTCTTTTTT TTTGAGGCC AAAGTATCAT 2520
1151
1152 GTGTAACAAC CGGTGTTCAA AAACGCGCCC GCCTGGCCGT TTAATCGCCC GATTAAATGA 2580
1153
1154 TGATCGGAAG GCTGCCATGG CGAGGCGGAG GTAATCAGTG GTTCTAGGCG CTGAAACTAG 2640
1155
1156 AAAACCTTCA AAAATCGAAA TTTTAAGAGC TAAATCGGTG TTTATCTCAT GAATCTATTA 2700
1157
1158 TATTTAGTTG AACTCACAA GAATCGGTTG TAAAACTAT GAAATCGTGC AAAAAAATG 2760
1159
1160 AAGAACAAAA TATTCTCAGA TCTGGAAC ACAGAGAAGA GGTGAAGAT GAGGGTAAAA 2820
1161
1162 TCGTATTTTG TCATTCATTA AACTAAAATC AAAAAAAAT GATGCAAAAT TCAATGATAA 2880
1163
1164 TAACTCGAAC TCGCAACCAT ATGCATCTTT AGACTGCGAC ACGGACCACT AGACTAAGCA 2940
1165
1166 ATTTTAATGT TTATTCATCA CAGACCTAAT ATATGTCTAA AACTAGGCGC CGAGTACGCC 3000

Raw Sequence Listing
Patent Application US/07/762,762

05/21/92
15:17:26

```
1167
1168 CCGCTTAATC CCGAGTTTTT GTTAGCTCGC TAGACCCAGG GTCACCGCCC GACTAACGAG 3060
1169
1170 TAGCGTAATT CTGAACTGGG GTAACAACAT AGAGAACATC GCCGACCCTT CCCTGCCGAT 3120
1171
1172 GATGCCGCCT CCGATGAACT TCCTGTAACG CCTTCAGTTT CCATTGATTT TCCCCTTTAA 3180
1173
1174 TCTGATCAGT TCCATGTTTT ATCCAACTCA TCCCACTCCG TAGCATTTAA TCGATCTCAT 3240
1175
1176 CATTTACATA CATAACCAGT AGGAGGTCTC ATATAAATTT GAACGTTTCC AGCGATGAAC 3300
1177
1178 AGTGCCAATC TCTGCGAAAT CCATTTCTCT AAGCTCAGGG CTGGCGGCTG CAGCCCGGGG 3360
1179
1180 ATCCACTAGT TCTAGGCGGC CGCACC GCGG TGGAGCTCCA ATTCGCCCTA TAGTGAGTCG 3420
1181
1182 TATTACGCGC GCTCACTGGC 3440
1183
1184
```

(2) INFORMATION FOR SEQ ID NO:28:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 3898 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

```
1201
1202 CTCGAGAGCT GAAGGATTTT TTGTTAGAGA TTCAACGACA GATGGACCCT TCCTCCACTA 60
1203
1204 GGCAACTGCA AGAACCTAAC AATGCAAATA TCACTCCTCC TCAGCCTTCA AGGAGCGTTA 120
1205
1206 ATAGGACTGG AACAAGCGGT CAAGTGAGTA AATTTTCCTT CCAAGATAGA TCTCTATGGT 180
1207
1208 TCGGTTTCATG AAGTTTGTGG TTAAATTGTG TAGCAACAGG ATAGTGCAAG TGAGAATAGA 240
1209
1210 GTTCGACCTC ATCTACCTAC CCCGGAACCT CTGAATGTAT CCCCATTGAA GAAGAAGAGG 300
1211
1212 GCAAATCCTG CACCCAGAAG GATAAAGAAA TTTTGGACGC CTGAAGAAGT GGCAGTTCTG 360
1213
1214 AGGGAAGGAG TAAAAGAGTA TGTCTACTAC TACTACTCTA TAATCAAGTT TCAAGAAGCT 420
1215
1216 GAGCTTGGCT CTCACCTTAT ATGTTTGATG TTGTTGTGCA GGTATGGTAA ATCATGGAAA 480
1217
1218 GAGATAAAGA ATGCAAACCC TGAAGTATTG GCAGAGAGGA CTGAGGTGAG AGAGCATGTC 540
1219
```

Patent Application US/07/762,762

1220 ACTTTTGTGT TACTCATCTG AATTATCTTA TATGCGAATT GTAAGTGGTA CTAAAAGGTT 600
1221
1222 TGTAACCTTTT GGTAGGTGGA TTTGAAGGAT AAATGGAGGA ACTTGCTTCG GTAGCGGTAA 660
1223
1224 CAAGTTTTAT ATTGCTATGA AGTTTTTTTG CCTGCGTGAC GTATCAGCAG CTGTGGAGAA 720
1225
1226 GATGGTATTA GAAAGGGTCT TTTACATTT TGTGTTGTGA CAAATATTAA TTCGGCCGGT 780
1227
1228 ATGGTTTGGT TAAGACTTGT TGAGAGACGT GTGGGGTTTT TTGATGTATA ATTAGTCTGT 840
1229
1230 GTTTAGAACG AAACAAGACT TGTTCGTAT GCTTTTTTTA ACTTGAGGGG GTTTGTTGTT 900
1231
1232 GTTAGTTAGG AACTTGACTT TGTCTCTTTC TCTCAAGATC TGATTGGTAA GGTCTGGGTG 960
1233
1234 GTAGTACTGT TTGGTTTAAT TTGTTTTGAC TATTGAGTCA CTGTGGCCCA TTGACTTTAA 1020
1235
1236 ATTAGGCTGG TATATTTTTT GGTTTAAAAC CGGTCTGAGA TAGTGCAATT TCGATTCAGT 1080
1237
1238 CAATTTTAAA TTCTTCAAGG TAATGGGCTG AATACTTGTA TAGTTTAAAG ACTTAACAGG 1140
1239
1240 CCTTAAAAGG CCCATGTTAT CATAAAACGT CATTGTTTAG AGTGCACCAA GCTTATAAAA 1200
1241
1242 TGTAGCCAGG CCTTAAAAGA CTTAACAGGC CTTAAAAGAC TTAACATTCC TTAAAAGGCC 1260
1243
1244 CATGTTATCA TAAAACGTCA TCGTTTTGAG TGCACCAAGC TAAATGTAGC CAGGCCTTAA 1320
1245
1246 AAGACTTAAC AGGCCTTAAA AGGCCCATGT TATCATAAAA CGCCGTCGTT TTGAGTGCAC 1380
1247
1248 CAAGCTTATA AATGTAGCCA GCTACCTCGG GACATCACGC TCTTTGTACA CTCCGCCATC 1440
1249
1250 TCTCTCTCTC TCGAGCAGAT CTCTCTCGGG AATATCGACA ATGTCGACCA CTTTCTGCTC 1500
1251
1252 TTCCGCTCTC ATGCAAGCCA CTTCTCTGGT AATCTCATCT CCTTCTTG TG TTCCAGATC 1560
1253
1254 GCTCTGATCA TACTTTCTTT TAGATCATTT GCCTCTGATC TGTGCTTGA TGTGTTTAA 1620
1255
1256 CTCTCCACGC ATGTTTGATT ATGTTGAGAA TTAGAAAAAA AATGTTAGCT TTACGAATCT 1680
1257
1258 TTAGTGATCA TTTCAATTGG ATTTGCAATC TTGTGTGACA TTTGAGGCTT GTGTAGATTT 1740
1259
1260 CGATCTGTAT TCATTTTGAA TCACAGCTAT AATAGTCATT TGAGTAGTAG TGTGTTTAA 1800
1261
1262 TGAACATGTT TTGTTGTATT GATGGAACAA ACAGGCAGCA ACAACGAGGA TTAGTTTCCA 1860
1263
1264 GAAGCCAGCT TTGGTTTCAA CGACTAATCT CTCCTTCAAC CTCCGCCGTT CAATCCCCAC 1920
1265
1266 TCGTTTCTCA ATCTCCTGCG CGGTATGTTT TCATTCTCAG CATTTATTTT GAGCTTGCTT 1980
1267
1268 GTCATGGTAC TCTCTCTAAT TGTCTATTTG GTTTATTAGG CCAAACCAGA GACGGTTGAG 2040
1269
1270 AAAGTGTCTA AGATAGTTAA GAAGCAGCTA TCACTCAAAG ACGACCAAAA GGTTCGTTGCG 2100
1271
1272 GAGACCAAGT TTGCTGATCT TGGAGCAGAT TCTCTCGACA CTGTAAGTCA TCAATCATTC 2160

Raw Sequence Listing
Patent Application US/07/762,76205/21/92
15:17:41

1273
1274 TCTTATGTGA ATAAAGAGAA CTTGAAGAGT TTGTTTTTAA CATATTAACT GAGTGTTTTG 2220
1275
1276 CATGCAGGTT GAGATAGTGA TGGGTTTAGA GGAAGAGTTT GATATCGAAA TGGCTGAAGA 2280
1277
1278 GAAAGCTCAG AAGATTGCTA CTGTGGAGGA AGCTGCTGAA CTCATTGAAG AGCTCGTTCA 2340
1279
1280 ACTTAAGAAG TAATTTTAGT ATTAAGAGCA GCCAAGGCTT TGTGTTGGTTT GTTGTTTTCA 2400
1281
1282 TAATCTTCCT GTCATTTTCT TTTTCTTTAA TGTGTCAAGC GACTCTGTTG GTTTAAAGTA 2460
1283
1284 GTATCTGTTT GCCATGGATC TCTCTCTATT TGTCGACTGA AAACCTTTTG TTTACACATG 2520
1285
1286 AAAGCTTGTT CTGTCTCTTT CTAAATCGA AATGCCAAAT GCGAGATTAG GGAATCTTGT 2580
1287
1288 ATTAACACAT ACATAAGTCA AAGAGTAGGC CCTAAGATGA CAATTTATAA ACAATCCTAT 2640
1289
1290 TCACATTGTA TATACAGGTT ATGATTATTC CCAATCAGCG TCAAAGAATC CAGCATCTTT 2700
1291
1292 CATCTCTGAA TAGTAGACAT TCTCCAAGTT CACATCTTCC TCCTGCACCA AAAACCAGTA 2760
1293
1294 CTAAATCATG AACATTGCAA TAATCACATG CCTAGGCGAG AGTTTTGGTG ATGTGGTGTT 2820
1295
1296 AGTGATAGTG ATACTGATGG TGCTAGAGCG GTTAAGAAGG ATTAACCTGG AAGAAGTCTG 2880
1297
1298 CAAGGAAAGT AACATAGAGA AGAGGAAGAT AGGAGTGGTA ACAAACACTT GTGATCCCAT 2940
1299
1300 ACAGCCTCCC AGCATTTTTT AAATGTTATT TCCTTACATA AAGAAACAAG AGAAGTCTGA 3000
1301
1302 CTAGATGATA TTTATATAGG ATAAGTGTTT TACCATAAGC CAAAGTGAGC GCCGTTTGCA 3060
1303
1304 AGAGCTAACC AGACAGTACA CGTTTGCCAT ATATCTCATC AACATGATCT GAAAAGTAAC 3120
1305
1306 ATATCACAGT TAATGAACAC AATGGTTACC TTGAGAAGCA AATCAAGACC TATAACAAGC 3180
1307
1308 CCAGAGATGA GGAAAGTCCG TGTCAACGCT TCACCGCCAT TCGCGTAGTT TCCTTGGAAG 3240
1309
1310 ACAAAGGCCA CCAACCAAAC TTAATTCCAG AAACAACACT CCAAATGTTG TCAACAAAGT 3300
1311
1312 CAATAGATTG CAACTACTT CGTTACAGGG TTGTATAGAT AATATAATAG AATAGTGGGA 3360
1313
1314 AGATAGTATA AATAAAATAA ATAAAAGATC CTATCGGTAA ATAGTTTATA ATATCGGGGG 3420
1315
1316 CGTATATAAA GTATAAAGA AACTCTTCTC CAATCCGACC GTTGAAAATC ACTCTCAATC 3480
1317
1318 TCTGGCGTAA CGACCGGATC GTTCGCGCGT AATTTTCGCT GCTATAAATA GAAACTTTCC 3540
1319
1320 TCTTCTGTTT CTCGATCAAA ATTTTTTTTT GGAAAAATTA AGTTTGAATC TATCGTAGAT 3600
1321
1322 GCTGTGACAA AAAAAAATTG TTTTATCGAA GATGAGAAAC ATGAGGCCTG TTCATGCAAG 3660
1323
1324 GAACCAGACC ACGGATCCAT CTTGCGCGAT GATGACGTCT CCTCTGATGA ATCGTCACGC 3720
1325

Patent Application US/07/762,762

1326 ACGGACAGGA TCCAACGCTG GACCAGCATC TAACGCCAAG AAAGCACAGA CGAAAGCAGC 3780
1327
1328 AGCTCAGAGA CTCGCGGCTG TGATGTCGAA CCAAACAGGC GACGATGAAG ACAGTGATGA 3840
1329
1330 TGACCTTTCC TTGACTACA ACGCTGTCGG AAGCATTGGT CTCGCTGCCG GAAGATCT 3898
1331

(2) INFORMATION FOR SEQ ID NO:29:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 4325 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

1348
1349
1350 CTCGAGGCAG TCACTAACAT GAAGTTTGAC GAGGAGCCCA ACTATGGGAA GCTTATTTCT 60
1351
1352 CTTTTCGATA CTCTAATTGA GCCGTGCGCT CTATCTAGAC CAATTAGAAT TGATGGAGCT 120
1353
1354 CTAAAGGTTG CTGGCTGTTT TCTTGTTTAT ATGATTAACT TCTAAACTTG TGTATAAATA 180
1355
1356 TTCTCTGAAA GTGCTTCTTT TGGCATATGT AGGTGTTGGCA AAAACGAGGA AGATTGCTTC 240
1357
1358 TCAATTTGGA AGAGGATGAA CAGCCGAAGA AGAAAATAAG AATAGGCAGT CCTGCTACTC 300
1359
1360 AATGGATCTC AGTCTATAAC GGTCTGTCGTC CCATGAAACA GAGGTAAAAC ATTTTTTTGCA 360
1361
1362 TATACACTTT GAAAGTTTCT CACTAACTGT GTAATCTTTT GGTAGATATC ACTACAATGT 420
1363
1364 CGGAGAGACA ANGGCTGSNC ANCATATACA AAAGGGAAAT GAAGATGGCC TTTTGATTAG 480
1365
1366 CTGTGTAGCA TCAGCAGCTA ATCTCTGGGC TCTCATCATG GATGCTGGAA CTGGATTAC 540
1367
1368 TTCTCAAGTT TATGAGTTGT CACCGGTCTT CCTACACAAG GTAATAATCA GTTGAAGCAA 600
1369
1370 TTAAGAATCA ATTTGATTTG TAGTAAACTA AGAAGAACTT ACCTTATGTT TTCCCCGCAG 660
1371
1372 GACTGGATTA TGAACAATG GGAAAAGAAC TACTATATAA GCTCCATAGC TGGTTCAGAT 720
1373
1374 AACGGGAGCT CTTTAGTTGT TATGTCAAAA GGTAGTGTG TAGTGAATAA TAAACTTATA 780
1375
1376 CCACAAAGTC TTCATTGACT TATTTATATA CTTGTTGTGA ATTGCTAGGA ACTACTTATT 840
1377
1378 CTCAGCAGTC ATACAAAGTG AGTGACTCAT TTCCGTTCAA GTGGATAAAT AAGAAATGGA 900

Patent Application US/07/762,762

1379
1380 AAGAAGATTT TCATGTAACC TCCATGACAA CTGCTGGTAA TCGTTGGGGT GTGGTAATGT 960
1381
1382 CGAGGAACTC TGGCTTCTCT GATCAGGTAG GTTTTTGTCT CTTATTGTCT GGTGTTTTTA 1020
1383
1384 TTTTCCCCTG ATAGTCTAAT ATGATAAACT CTGCGTTGTG AAAGGTGGTG GAGCTTGACT 1080
1385
1386 TTTTGTACCC AAGCGATGGG ATACATAGGA GGTGGGAGAA TGGGTATAGA ATAACATCAA 1140
1387
1388 TGGCAGCAAC TGC GGATCAA GCAGCTTTCA TATTAAGCAT ACCAAAGCGT AAGATGGTGG 1200
1389
1390 ATGAAACTCA AGAGACTCTC CGCACCACCG CCTTCCAAG TACTCATGTC AAGGTTGGTT 1260
1391
1392 TCTTTAGCTT TGAACACAGA TTTGGATCTT TTTGTTTTGT TTCCATATAC TTAGGACCTG 1320
1393
1394 AGAGCTTTTG GTTGATTTTT TTTTCAGGAC AAATGGGCGA AGAATCTGTA CATTGCATCA 1380
1395
1396 ATATGCTATG GCAGGACAGT GTGCTGATAC ACACCTAAGC ATCATGTGGA AAGCCAAAGA 1440
1397
1398 CAATTGGAGC GAGACTCAGG GTCGTCATAA TACCAATCAA AGACGTAAAA CCAGACGCAA 1500
1399
1400 CCTCTTTGGT TGAATGTAAT GAAAGGGATG TGTCTTGGTA TGTATGTACG AATAACAAAA 1560
1401
1402 GAGAAGATGG AATTAGTAGT AGAAATATTT GGGAGCTTTT TAAGCCCTTC AAGTGTGCTT 1620
1403
1404 TTTATCTTAT TGATATCATC CATTTGCGTT GTTAAATGCG TCTCTAGATA TGTTCCTATA 1680
1405
1406 TCTTTCTCAG TGTCTGATAA GTGAAATGTG AGAAAACCAT ACCAAACCAA AATATTCAAA 1740
1407
1408 TCTTATTTTT AATAATGTTG AATCACTCGG AGTTGCCACC TTCTGTGCCA ATTGTGCTGA 1800
1409
1410 ATCTATCACA CTAGAAAAAA ACATTTCTTC AAGGTAATGA CTTGTGGACT ATGTTCTGAA 1860
1411
1412 TTCTCATTAA GTTTTTATTT TCTGAAGTTT AAGTTTTTAC CTTCTGTTTT GAAATATATC 1920
1413
1414 GTTCATAAGA TGTCACGCCA GGACATGAGC TACACATCGC ACATAGCATG CAGATCAGGA 1980
1415
1416 CGATTTGTCA CTCACTTCAA ACACCTAAGA GCTTCTCTCT CACAGCGCAC ACACATATGC 2040
1417
1418 ATGCAATATT TACACGTGAT CGCCATGCAA ATCTCCATTC TCACCTATAA ATTAGAGCCT 2100
1419
1420 CGGCTTCACT CTTTACTCAA ACCAAAACCTC ATCACTACAG AACATACACA AATGGCGAAC 2160
1421
1422 AAGCTCTTCC TCGTCTCGGC AACTCTCGCC TTGTTCTTCC TTCTCACCAA TGCCTCCGTC 2220
1423
1424 TACAGGACGG TTGTGGAAGT CGACGAAGAT GATGCCACAA ATCCAGCCGG CCCATTTAGG 2280
1425
1426 ATTCCAAAAT GTAGGAAGGA GTTTCAGCAA GCACAACACC TGAAAGCTTG CCAACAATGG 2340
1427
1428 CTCCACAAGC AGGCAATGCA GTCCGGTAGT GGTCCAAGCT GGACCCTCGA TGGTGAGTTT 2400
1429
1430 GATTTTGAAG ACGACGTGGA GAACCAACAA CAGGGCCCCG AGCAGAGGCC ACCGCTGCTC 2460
1431

Patent Application US/07/762,762

1432 CAGCAGTGCT GCAACGAGCT CCACCAGGAA GAGCCACTTT GCGTTTGCCC AACCTTGAAA 2520
1433
1434 GGAGCATCCA AAGCCGTTAA ACAACAGATT CGACAACAAC AGGGACAACA AATGCAGGGA 2580
1435
1436 CAGCAGATGC AGCAAGTGAT TAGCCGTATC TACCAGACCG CTACGCACTT ACCTAGAGCT 2640
1437
1438 TGCAACATCA GGCAAGTTAG CATTTGCCCC TTCCAGAAGA CCATGCCTGG GCGCGGCTTC 2700
1439
1440 TACTAGATTC CAAACGAATA TCCTCGAGAG TGTGTATACC ACGGTGATAT GAGTGTGGTT 2760
1441
1442 GTTGATGTAT GTTAACACTA CATAGTCATG GTGTGTGTTT CATAAATAAT GTACTAATGT 2820
1443
1444 AATAAGAACT ACTCCGTAGA CGGTAATAAA AGAGAAGTTT TTTTTTTTAC TCTTGCTACT 2880
1445
1446 TTCCTATAAA GTGATGATTA ACAACAGATA CACCAAAAAG AAAACAATTA ATCTATATTC 2940
1447
1448 ACAATGAAGC AGTACTAGTC TATTGAACAT GTCAGATTTT CTTTTTCTAA ATGTCTAATT 3000
1449
1450 AAGCCTTCAA GGCTAGTGAT GATAAAAGAT CATCCAATGG GATCCAACAA AGACTCAAAT 3060
1451
1452 CTGGTTTTGA TCAGATACTT CAAAACATTT TTTGTATTCA TTAAATTATG CAAGTGTTCT 3120
1453
1454 TTTATTTGGT GAAGACTCTT TAGAAGCAAA GAACGACAAG CAGTAATAAA AAAAACAAAAG 3180
1455
1456 TTCAGTTTTA AGATTTGTGA TTGACTTATT GTCATTTGAA AAATATAGTA TGATATTAAT 3240
1457
1458 ATAGTTTTAT TTATATAATG CTTGTCTATT CAAGATTTGA GAACATTAAT ATGATACTGT 3300
1459
1460 CCACATATCC AATATATTAA GTTTCATTTT TGTTCAAAACA TATGATAAGA TGGTCAAATG 3360
1461
1462 ATTATGAGTT TTGTTATTTA CCTGAAGAAA AGATAAGTGA GCTTCGAGTT TCTGAAGGGT 3420
1463
1464 ACGTGATCTT CATTTCTTGG CTAAAAGCGA ATATGACATC ACCTAGAGAA AGCCGATAAT 3480
1465
1466 AGTAAACTCT GTTCTTGGTT TTTGGTTTAA TCAAACCGAA CCGGTAGCTG AGTGTCAAGT 3540
1467
1468 CAGCAACAT CGCAACCAT ATGTCAATTC GTTAGATTCC CGGTTTAAAGT TGTAAACCGG 3600
1469
1470 TATTTTATTT GGTGAAAACC CTAGAAGCCA GCCANCCTTT TTAATCTAAT TTTTGCAAAC 3660
1471
1472 GAGAAGTCAC CACACCTCTC CACTAAAACC CTGAACCTTA CTGAGAGAAG CAGAGNCANN 3720
1473
1474 AAAGAACAAA TAAAACCCGA AGATGAGACC ACCACGTGCG GCGGGACGTT CAGGGGACGG 3780
1475
1476 GGAGGAAGAG AATGRCGGCG GNSNTTGGT GCGGCGGGCG GACGTTTTGG TGGCGGCGGT 3840
1477
1478 GGACGTTTTG GTGGCGGCGG TGGACCTTTG GTGGTGGATA TCGTGACGAA GGACCTCCCA 3900
1479
1480 GTGAAGTCAT TGGTTCGTTT ACTCTTTTCT TAGTCGAATC TTATTCTTGC TCTGCTCGTT 3960
1481
1482 GTTTTACCGA TAAAGCTTAA GACTTTATTG ATAAAGTTCT CAGCTTTGAA TGTGAATGAA 4020
1483
1484 CTGTTTCCTG CTTATTAGTG TTCCTTTGTT TTGAGTTGAA TCACTGTCTT AGCACTTTTG 4080

Raw Sequence Listing
Patent Application US/07/762,762

05/21/92
15:18:10

1485
1486 TTAGATTCAT CTTTGTGTTT AAGTTAAAAG GTAGAAACTT TGTGACTTGT CTCCGTTATG 4140
1487
1488 ACAAGGTTAA CTTTGTGTTT TATAACAGAA GTTGCGACCT TTCTCCATGC TTGTGAGGGT 4200
1489
1490 GATGCTGTGG ACCAAGCTCT CTCAGGCGAA GATCCCTTAC TTCAATGCCC CAATCTACTT 4260
1491
1492 GGAAAACAAG ACACAGATTG GGAAAGTTGA TGAGATCCAA GCTTGGGCTG CAGGTGCGACG 4320
1493
1494 AATTC 4325
1495

1496 (2) INFORMATION FOR SEQ ID NO:30:

1497
1498 (i) SEQUENCE CHARACTERISTICS:

1499
1500 (A) LENGTH: 30 base pairs

1501
1502 (B) TYPE: nucleic acid

1503
1504 (C) STRANDEDNESS: single

1505
1506 (D) TOPOLOGY: linear

1507
1508 (ii) MOLECULE TYPE: other nucleic acid

1509
1510 (A) DESCRIPTION: synthetic oligonucleotide

1511
1512 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

1513
1514 CGGATCCACT GCAGTCTAGA GGGCCCGGGA 30
1515
1516

1517 (2) INFORMATION FOR SEQ ID NO:31:

1518
1519 (i) SEQUENCE CHARACTERISTICS:

1520
1521 (A) LENGTH: 38 base pairs

1522
1523 (B) TYPE: nucleic acid

1524
1525 (C) STRANDEDNESS: single

1526
1527 (D) TOPOLOGY: linear

1528
1529 (ii) MOLECULE TYPE: other nucleic acid

1530
1531 (A) DESCRIPTION: synthetic oligonucleotide

1532
1533 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

1534
1535 AATTTCCCGG GCCCTCTAGA CTGCAGTGGA TCCGAGCT 38
1536
1537

Raw Sequence Listing
Patent Application US/07/762,762

05/21/92
15:18:18

1538 (2) INFORMATION FOR SEQ ID NO:32:

1539

1540 (i) SEQUENCE CHARACTERISTICS:

1541

1542 (A) LENGTH: 50 base pairs

1543

1544 (B) TYPE: nucleic acid

1545

1546 (C) STRANDEDNESS: single

1547

1548 (D) TOPOLOGY: linear

1549

1550 (ii) MOLECULE TYPE: other nucleic acid

1551

1552 (A) DESCRIPTION: synthetic oligonucleotide

1553

1554 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

1555

1556 GTAAGTAGGT AGGGCTTCCT CTGTAATCAT ATCTCCAACC AAAACAACAA 50

1557

1558

1559

1560 (2) INFORMATION FOR SEQ ID NO:33:

1561

1562 (i) SEQUENCE CHARACTERISTICS:

1563

1564 (A) LENGTH: 39 base pairs

1565

1566 (B) TYPE: nucleic acid

1567

1568 (C) STRANDEDNESS: single

1569

1570 (D) TOPOLOGY: linear

1571

1572 (ii) MOLECULE TYPE: other nucleic acid

1573

1574 (A) DESCRIPTION: synthetic oligonucleotide

1575

1576 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

1577

1578 CTTAAGAAGT AACCCGGGCT GCAGTTT TAG TATTAAGAG 39

1579

1580

1581 (2) INFORMATION FOR SEQ ID NO:34:

1582

1583 (i) SEQUENCE CHARACTERISTICS:

1584

1585 (A) LENGTH: 43 base pairs

1586

1587 (B) TYPE: nucleic acid

1588

1589 (C) STRANDEDNESS: single

1590

Raw Sequence Listing
Patent Application US/07/762,762

05/21/92
15:18:25

1591 (D) TOPOLOGY: linear
1592
1593 (ii) MOLECULE TYPE: other nucleic acid
1594
1595 (A) DESCRIPTION: synthetic oligonucleotide
1596
1597 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:
1598
1599 GGAATTCGTC GACAGATCTC TGCAGCTCGA GGGATCCAAG CTT 43
1600
1601
1602 (2) INFORMATION FOR SEQ ID NO:35:
1603
1604 (i) SEQUENCE CHARACTERISTICS:
1605
1606 (A) LENGTH: 48 base pairs
1607
1608 (B) TYPE: nucleic acid
1609
1610 (C) STRANDEDNESS: single
1611
1612 (D) TOPOLOGY: linear
1613
1614 (ii) MOLECULE TYPE: other nucleic acid
1615
1616 (A) DESCRIPTION: synthetic oligonucleotide
1617
1618 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:
1619
1620 CCATTTTGA TCTTCCTCGA GCCCGGGCTG CAGTTCTTCT TCTTCTTG 48
1621
1622
1623 (2) INFORMATION FOR SEQ ID NO:36:
1624
1625 (i) SEQUENCE CHARACTERISTICS:
1626
1627 (A) LENGTH: 48 base pairs
1628
1629 (B) TYPE: nucleic acid
1630
1631 (C) STRANDEDNESS: single
1632
1633 (D) TOPOLOGY: linear
1634
1635 (ii) MOLECULE TYPE: other nucleic acid
1636
1637 (A) DESCRIPTION: synthetic oligonucleotide
1638
1639 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:
1640
1641 GCTCGTTTTT TTTTCTCTG CAGCCGGGC TCGAGTCACA GCTTCACC 48
1642
1643

Raw Sequence Listing
Patent Application US/07/762,762

05/21/92
15:18:32

1644 (2) INFORMATION FOR SEQ ID NO:37:

1645

1646 (i) SEQUENCE CHARACTERISTICS:

1647

1648 (A) LENGTH: 44 base pairs

1649

1650 (B) TYPE: nucleic acid

1651

1652 (C) STRANDEDNESS: single

1653

1654 (D) TOPOLOGY: linear

1655

1656 (ii) MOLECULE TYPE: other nucleic acid

1657

1658 (A) DESCRIPTION: synthetic oligonucleotide

1659

1660 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:

1661

1662 ACTGACTGCA GCCCCGGGCTC GAGGAAGATC AAAAATGGCT CTTC 44

1663

1664

1665

1666 (2) INFORMATION FOR SEQ ID NO:38:

1667

1668 (i) SEQUENCE CHARACTERISTICS:

1669

1670 (A) LENGTH: 43 base pairs

1671

1672 (B) TYPE: nucleic acid

1673

1674 (C) STRANDEDNESS: single

1675

1676 (D) TOPOLOGY: linear

1677

1678 (ii) MOLECULE TYPE: other nucleic acid

1679

1680 (A) DESCRIPTION: synthetic oligonucleotide

1681

1682 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:

1683

1684 GAGTAGTGAA CTTCATGGAT CCTCGAGGTC TTGAAAACCT AGA 43

1685

1686

1687 (2) INFORMATION FOR SEQ ID NO:39:

1688

1689 (i) SEQUENCE CHARACTERISTICS:

1690

1691 (A) LENGTH: 44 base pairs

1692

1693 (B) TYPE: nucleic acid

1694

1695 (C) STRANDEDNESS: single

1696

Patent Application US/07/762,762

1697 (D) TOPOLOGY: linear
1698
1699 (ii) MOLECULE TYPE: other nucleic acid
1700
1701 (A) DESCRIPTION: synthetic oligonucleotide
1702
1703 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:
1704
1705 CAATGTCTTG AGAGATCCCG GGATCCTTAA CAACTAGGAA AAGG 44
1706
1707
1708 (2) INFORMATION FOR SEQ ID NO:40:
1709
1710 (i) SEQUENCE CHARACTERISTICS:
1711
1712 (A) LENGTH: 24 base pairs
1713
1714 (B) TYPE: nucleic acid
1715
1716 (C) STRANDEDNESS: single
1717
1718 (D) TOPOLOGY: linear
1719
1720 (ii) MOLECULE TYPE: other nucleic acid
1721
1722 (A) DESCRIPTION: synthetic oligonucleotide
1723
1724 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:
1725
1726 GTAAGACACG ACTTATCGCC ACTG 24
1727
1728
1729 (2) INFORMATION FOR SEQ ID NO:41:
1730
1731 (i) SEQUENCE CHARACTERISTICS:
1732
1733 (A) LENGTH: 43 base pairs
1734
1735 (B) TYPE: nucleic acid
1736
1737 (C) STRANDEDNESS: single
1738
1739 (D) TOPOLOGY: linear
1740
1741 (ii) MOLECULE TYPE: other nucleic acid
1742
1743 (A) DESCRIPTION: synthetic oligonucleotide
1744
1745 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:
1746
1747 GGAATTCGTC GACAGATCTC TGCAGCTCGA GGGATCCAAG CTT 43
1748
1749

Raw Sequence Listing
Patent Application US/07/762,762

05/21/92
15:18:47

1750 (2) INFORMATION FOR SEQ ID NO:42:

1751

1752 (i) SEQUENCE CHARACTERISTICS:

1753

1754 (A) LENGTH: 32 base pairs

1755

1756 (B) TYPE: nucleic acid

1757

1758 (C) STRANDEDNESS: single

1759

1760 (D) TOPOLOGY: linear

1761

1762 (ii) MOLECULE TYPE: other nucleic acid

1763

1764 (A) DESCRIPTION: synthetic oligonucleotide

1765

1766 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:42:

1767

1768 GCTTGTTTCGC CATGGATATC TTCTGTATGT TC 32

1769

1770 (2) INFORMATION FOR SEQ ID NO:43:

1771

1772

1773 (i) SEQUENCE CHARACTERISTICS:

1774

1775 (A) LENGTH: 143 base pairs

1776

1777 (B) TYPE: nucleic acid

1778

1779 (C) STRANDEDNESS: double

1780

1781 (D) TOPOLOGY: linear

1782

1783 (ii) MOLECULE TYPE: cDNA to mRNA

1784

1785

1786 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:43:

1787

1788 GAT GCC AAA ANG CCT CAC ATG CCT CCT AGA GAA GCT CAT GTG CAA AAG 48

1789 Asp Ala Lys Xaa Pro His MET Pro Pro Arg Glu Ala His Val Gln Lys

1790 1 5 10 15

1791

1792 ACC CAT TCA ATK CCG CCT CAA AAG ATT GAG ATT TTC AAA TCC TTG GAG 96

1793 Thr His Ser Xaa Pro Pro Gln Lys Ile Glu Ile Phe Lys Ser Leu Glu

1794 20 25 30

1795

1796 GGT TGG GCT GAG GAG AAT GTC TTG GTG CAT CTT AAA CCT GTG GAG AA 143

1797 Gly Trp Ala Glu Glu Asn Val Leu Val His Leu Lys Pro Val Glu

1798 35 40 45

1799

1800

1801

PAGE: 1

SEQUENCE VERIFICATION REPORT
PATENT APPLICATION US/07/762,762

DATE: 05/21/92
TIME: 15:18:55

LINE ERROR

ORIGINAL TEXT

39 Wrong application Serial Number
51 Unknown or Misplaced Identifier

(A) APPLICATION NUMBER: 07/762,762
(C) CLASSIFICATION:

PAGE: 1

SEQUENCE MISSING ITEM REPORT
PATENT APPLICATION US/07/762,762

DATE: 05/21/92
TIME: 15:18:55

MANDATORY IDENTIFIER THAT WAS NOT FOUND

PAGE: 1

SEQUENCE CORRECTION REPORT
PATENT APPLICATION US/07/762,762

DATE: 05/21/92
TIME: 15:18:55

LINE ORIGINAL TEXT

CORRECTED TEXT